


WHAT IS CLAIMED IS:

1.  A system for retrieval and processing of a data set from one or more data sources comprising:

5 a query structure assembly module for defining a query structure based upon a plurality of query assembly rules and a desired data set;

a syntax assembly module for defining at least one query language statement based upon the defined query structure;

a process optimization module for evaluating processing options based upon a database schema associated with the data source; and

10 whereby at least one query language statement may assembled and run against the data source(s) to return the desired data result set.

2. The system of claim 1, wherein the process optimization module includes a table selection module for evaluating the size of a selected set of tables from the database schema.

15 3. The system of claim 1, wherein the process optimization module includes an intermediate data processing selection module for evaluating the reusability of an intermediate data set in returning the defined data result set.

20 4. The system of claim 1, wherein the process optimization module includes an intermediate data processing method module for evaluating a plurality of methods for generating intermediate data sets within the data source(s).

5. The system of claim 1, wherein the process optimization module includes a join path selection module for evaluating the length of at least one join path used in returning the defined data result set.

25 6. The system of claim 1, wherein the query structure assembly module accesses a query assembly rule associated with a selected database management system.

7. The system of claim 1, wherein the syntax assembly module accesses a syntax description associated with a selected database management system.

8. The system of claim 1, wherein the process optimization module accesses a query assembly rule, a syntax description, or a syntax pattern associated with a selected database management system.

9. The system of claim 1, wherein the system is a component in an online analytical processing systems, a reporting system, a business intelligence system, or a data mining system.

10. A method of generating a query language statement to be run against one or more data sources, comprising the steps of:

generating a query structure based upon a database schema associated with the data source and a desired data result set;

generating query language syntax based upon the query structure for returning the desired data result set from the data source(s); and

evaluating a plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization.

11. The method of claim 10, wherein the step of evaluating the plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization includes evaluating the size of a plurality of sets of identified tables for returning the desired data result set.

12. The method of claim 10, wherein the step of evaluating the plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization includes evaluating the reusability of intermediate data sets.

13. The method of claim 10, wherein the step of evaluating the plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization includes evaluating a plurality of methods for generating intermediate data sets.

5 14. The method of claim 10, wherein the step of evaluating a plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization includes evaluating a plurality of join paths used in returning the desired data result set.

10 15. The method of claim 10, wherein the step of evaluating a plurality of query assembly rules, syntax descriptions, or syntax patterns for process optimization includes evaluating at least one query assembly rule, syntax description, or syntax pattern associated with a selected database management system.

16. The method of claim 10, wherein the step of generating a query structure includes evaluating at least one query assembly rule associated with a selected database management system.

15 17. The method of claim 10, wherein the step of generating query language syntax includes evaluating at least one syntax description or syntax pattern associated with a selected database management system.

18. The method of claim 10, wherein the method is implemented in an online analytical processing systems, a reporting system, a business intelligence system, or a data mining system.

20 19. A medium having a processor readable program code embodied therein for retrieving and processing data from one or more data sources comprising:

code for causing the processor to evaluate a plurality of sets of tables within the data source(s) for generating a desired data result set;

25 code for causing the processor to evaluate at least one intermediate data set for reusability in generating the desired data result set;

code for causing the processor to evaluate a plurality of methods for generating intermediate data sets for use in generating the desired data result set;

code for causing the processor to evaluate a plurality of join paths used for joining tables to return the desired data result set; and

5

code for causing the processor to assemble at least one query language statement based upon the evaluations of the plurality of sets of tables, the at least one intermediate data set, the plurality of methods for generating intermediate data sets, and the plurality of join paths.

20. The medium of claim 19, further comprising code for causing the processor to evaluate at least one query assembly rule associated with a selected database management system.

10

093349-061504
T06T90-664E3350